Park Service Experience with Developing Monitoring and QA/QC Guidance Consistent with that of Other Federal Agencies and States

Roy J. Irwin, National Park Service, Water Resources Division, Fort Collins, CO

Biographical Sketch:

After receiving a bachelor's degree from the University of Kansas, Roy obtained his Ph.D. in Biology from Tulane University in 1970. He was an Assistant Professor at the University of New Orleans, and then took a series of positions with EPA in Washington, D.C. He headed up the FWS Power Plant Team in Ann Arbor, MI. Roy also served as a Contaminants Specialist for the FWS in Texas, then joined the Park Service in 1992.Current interests include study design, QA/QC, and statistical issues related to monitoring of contaminants and biological parameters in freshwater and marine aquatic environments.

Abstract:

In developing guidance for how Park networks should write detailed study plans including Quality Assurance Project Plans for long-term vital signs monitoring, the NPS has reviewed the QA/QC and Data Quality Objective guidance of the U.S. Environmental Protection Agency (EPA) and the US Geological Survey (USGS) and some other federal agencies, as well as the credible data statutes of several states. Since the Park Service puts water quality monitoring data into the EPA STORET database, we have also reviewed the entry and metadata explanations for QC information in the new STORET. Some of the major differences in approach in terminology between USGS, EPA, and others, are reviewed. In spite of some differences and incompatibilities in the approaches of different groups, the NPS has been able to draft guidance that is reasonably compatible with most other widely used guidance packages. Our planning process steps are presented, as well as highlights of issues that had to be resolved to develop guidance reasonably consistent with that of other agencies.

Introduction:

The National Park Service (NPS) has a new long term monitoring program, still in an early design phase, called the "Park Vital Signs Monitoring" program. In developing guidance for the water quality portion of the program, the NPS wanted to come up with state-of-the-art project planning and Quality Assurance/Quality Control (QA/QC) guidance. An important NPS goal was that the guidance should be generally consistent with QA/QC guidance of the EPA, USGS, State Credible Data Statutes, and recent thinking relating to "Sound Science" and peer review.

Mark Twain said "Always do the right thing. This will gratify some people and astonish the rest." However, determining the right thing is not always that easy. Major agencies like EPA and USGS don't seem to have uniform ideas on what those "right things" are. Hopefully, the National Water Quality Monitoring Council (NWQMC) can help standardize some "right things" in federal water quality monitoring.

General Differences Between EPA and USGS Guidance:

Reviewing relevant guidance from other agencies revealed that EPA had a great many national, regional and program-specific guidance documents relating to study design and QA/QC, and that not all of them gave consistent guidance or even used consistent terminology. EPA's STORET database suggests one use terminology and options not particularly consistent with various EPA guidance documents. So when referring to EPA QA/QC guidance, the first question one might ask is "Which one?"

USGS guidance documents relevant to study design and QA/QC seemed to be fewer in number and harder for outsiders to find than those for EPA. USGS has standard methods for freshwater habitats posted on the Internet in the form of their field manual (http://water.usgs.gov/owq/FieldManual) and in the form of NAWQA protocols (http://water.usgs.gov/nawqa/protocols/doc_list.html). However, guidance for how to perform QA/QC or how to develop a Quality Assurance Project Plan (QAPP) is generally not available to those outside of USGS. The only national guidance we could find related to QA/QC was one for toxic chemicals, a NAWQA document posted at http://water.usgs.gov/nawqa/protocols/OFR97-223/ofr97-223.pdf, which lists frequencies of QC samples but not recommended performance standards or measurement quality objectives for things like precision and bias. The USGS Water Resources Division (WRD) regions are responsible for routine water quality monitoring all over the nation. The only USGS QA/QC guidance document we could find that related to their work, or to USGS performance goals for standard water quality parameters like pH and dissolved oxygen, was a QAPP authored by a Texas USGS WRD that was put on the Internet by the North Central Texas Council of Governments (http://www.dfwstormwater.com/FY01/PDFs/appendix_3_USGS.pdf).

Looking through the documents from EPA and USGS and talking to experts from both agencies, different approaches to QA/QC became apparent:

For example, the EPA guidance documents tend to suggest that one determine quantitative pre-project measurement quality data acceptance criteria or measurement quality objectives for data quality indicators like precision and bias. If data do not meet these criteria, the tendency is to throw the data out (or at least not use the data for critical decisions unless in a highly qualified manner), or recalibrate the instruments, or make other corrections until the measurement process does meet criteria before repeating the analyses. Many states and independent labs also evidently use this process. The reliance on pre-project data acceptance criteria for data quality indicators is consistent with Performance Based Measurement Systems (PBMS) needs to assess "comparability" of data generated by different methods, different labs, different operators, and different days. The emphasis on defining acceptance criteria for data quality indicators is also consistent with the need to determine if data quality objectives are appropriate for the monitoring purpose and the need to assess the quality of data from a large variety of labs. Unlike USGS, the EPA, NPS, and States use data from labs all over the country to generate data relevant to EPA and State regulatory processes. The NPS, many States, and many parts of EPA archive data in STORET.

There is not 100% uniformity in how QA/QC is handled either in EPA or in USGS, and thus we are using the phrase "tends to be" here rather than stronger language. However, there are some ways in which the typical USGS approach to QA/QC does seem to be fundamentally different than the typical approaches recommended by EPA. Many projects using EPA guidance for short-term or even one-time projects use the EPA recommended methods to help answer specific regulatory questions.

Although it does have some short-term synoptic projects, in general the USGS tends to have many longer term monitoring projects and sends most of is water samples to one lab (the NWQL in Denver). The USGS also tends to rely heavily on the QA benefits of uniform training of technicians. To a greater degree than many other groups, USGS tends to assess QA/QC over longer periods of time. This multi-year approach to assessing QA/QC typically involves looking at all results for a certain parameter and program for a set period of time, say 1997 to 2001. Those doing short-term or one-time projects using EPA or State methods and using multiple labs typically do not have this luxury. They have to decide if the data is acceptable for a given use or decision soon after it is collected.

Although USGS will not publish data that is clearly wrong (way outside of control limits or possible values, for example), and will try to correct observed control limit problems as they crop up, in general USGS does not tend to use the phrase "acceptance criteria." Accordingly, many groups in USGS seem to have fewer tendencies to throw out or not publish data than those that are trying to follow EPA suggestions. If the recovery of a difficult to analyze pesticide is only 20%, the USGS might report the value obtained with proper qualifiers rather than throwing the data out as not meeting a bias acceptance criterion. USGS publishes its data in its own database, one that has different QC descriptor defaults than does EPA's STORET. Again, although there are hints of differences in overall approaches between EPA and USGS, there are also similarities. Both agencies tend to recommend controlling the measurement process and both agencies tend to suggest annotating some data with explanatory QC notation code letters or numbers. Both say it is up to user to interpret data quality.

Differences in More Specific Terminology, Criteria, and Goals:

> Precision: Most EPA documents use the word precision. Some in USGS use that word too but others simply use the term variability. Common summary statistics used vary from Relative Percent Difference (RPD) and Relative Standard Deviation (RSD) (both recommended by EPA EMAP) to RPD-only (recommended by some groups in USGS) to Confidence Intervals (recommended in EPA STORET).
>
> Bias: Most newer EPA documents correctly distinguish between bias and accuracy. Some USGS and EPA documents wrongly make the two terms synonyms.

Accuracy and Uncertainty: Most USGS documents that I have been able to find are silent on how to quantify overall accuracy and uncertainty. Many EPA documents are also silent on the subject. Those that do address it vary on what they say about it. Some EPA documents use one of the equations also found in older versions of Standard Methods to quantify uncertainty in a measurement process. The Park Service has also found this equation [What the NPS calls Root Square Accuracy Error = the square root of (sample variance + net bias squared] to be helpful in quantifying minimum uncertainty in a measurement process. The latest version of Standard Methods suggests bias-adjusted confidence interval approaches. The NPS also recommends such approaches for estimating minimum measurement uncertainty in the mean value of the parameter being measured.

Error: The term "error" is not used or defined in most USGS documents I have read. The word error is used for hypothesis error testing rates (type I or type II) in some EPA DQO documents. Most agree that "error" has something to do with variability. ISO defines it as variability (in standard deviation units).

Standard Statistics Used: The recommendations are highly variable in EPA QA/QC guidance documents, though there is a tendency to recommend parametric statistics and hypothesis testing, especially in documents that explain the Data Quality Objective (DQO) process. Certain nonparametric methods are becoming so common in USGS documents that they seem to have become default standards in the USGS (L.M. Griffith, R.C. Ward, G.B. McBride, and J.C. Loftis. 2001. Data Analysis Considerations in Producing 'Comparable' Information for Water Quality Management Purposes. National Water Quality Monitoring Council Technical Report 01-01. White Paper of the National Water Quality Monitoring Council, Co-sponsored by USGS, Web: http://water.usgs.gov/wicp/acwi/monitoring/CouncilPrior6-Mar00.html).
.
Detection Limits: The method detection limit (MDL) as defined in 40 CFR Part 136 seems to be universal in most EPA and State settings for a semi-qualitative detection limit. USGS now uses a similar term (LT-MDL) based on long-term nonparametric statistics. Quantitative detection limit terminology and how far above a MDL the quantitative limit is, and how to use and/or annotate values between the two limits, varies tremendously in various EPA QA, STORET, and USGS documents.

Sensitivity Performance Criteria for Field Probes. So far we have not found much guidance in either agency, and manufacturers are of variable help.

Precision Performance Goals for Measurement Quality Criteria for Field Probes: For parameters like pH or dissolved oxygen, EPA's defaults for E-EMAP program are plus or minus 10% for either RPD or RSDs. USGS Texas WRD "QA Goals" are plus or minus 20% RPDs.

QC annotation codes: These vary between various EPA and USGS documents. The new STORET seems to have unique codes that are new (and different than past codes recommended by various agencies).

Where to measure in the stream and other details: This varies greatly between EPA's EMAP guidance, various other EPA documents, and the different USGS documents.

Reporting Bias-Adjusted Data: Most labs don't report bias-adjusted data, yet new STORET, by use of its codes and default choices, seems to hint that data should be adjusted for bias. Labs worry about being accused of malice if they adjust data and hope data users know whether to adjust the data or not. Many data users do not understand the subject. NPS has tried to explain the issue and to suggest properly framed and bias-adjusted confidence intervals as one option for addressing measurement accuracy or uncertainty.

The National Park Service Approach:

The Park Service recommends a planning process outline and detailed discussions of the topics discussed above, trying to be generally consistent with both EPA, USGS, modern "sound science" concepts, and up-to-date pre-project peer review recommendations of the Whitehouse and others. We have also incorporated the concept of honestly admitting uncertainty, a key concept in the new law requiring agencies to publish Data Quality Guidelines (Public Law 106-554, see http://www.epa.gov/oei/qualityguidelines/index.htm). Finally, the NPS has incorporated the relevant parts of EPA's DQO process into our monitoring guidance.

The NPS outline of things to consider in a detailed study plan that includes a Quality Assurance Project Plan is detailed at http://www.nature.nps.gov/im/monitor/wqPartB.doc.

The draft includes the following steps (peer review comments welcome, send them to Roy_Irwin@nps.gov):

GENERAL PROJECT PLANNING STEPS ARE ALL PART OF QUALITY ASSURANCE (QA):

Including the Preliminary steps:

1) GATHER AND SUMMARIZE INFORMATION NEEDED

2) IDENTIFY CONTEXT (REGULATORY OR GENERAL)

3) PLAN AS A TEAM (PRE-PROJECT PEER REVIEW)

QUALITY ASSURANCE/GENERAL PROJECT PLANNING STEPS:

4. PROBLEM STATEMENT/VALUE (S) TO BE PROTECTED

5. BACKGROUND

6. QUESTIONS TO BE ANSWERED/OBJECTIVES

7. IDENTIFICATION OF DECISIONS AND DECISION RULES (THRESHOLD RESULTS THAT WILL TRIGGER IDENTIFIED CONCLUSIONS OR ACTIONS)

8. IDENTIFICATION OF TARGET POPULATION, STUDY BOUNDARIES, & SAMPLE UNITS

9. WHAT WILL BE MEASURED?

10. RELATIONSHIP AND CONCEPTUAL MODELS

11. PROJECT MANAGEMENT BUDGET, STAFF QUALIFICATIONS, AND STAFF TRAINING

12. DATA MANAGEMENT, DATA HANDLING, REPORTING, AND ARCHIVING

13. OVERALL MONITORING DESIGN, STATISTICS, STATISTICAL POWER, AND DATA QUANTITY OBJECTIVES

14. LAB SELECTION

QC OBJECTIVES FOR DATA QUALITY INDICATORS

15. DATA REPRESENTATIVENESS

16. DATA COMPARABILITY/STANDARD OPERATING PROCEDURES (SOPs)/STANDARD PROTOCOLS

17. DETECTION LIMITS/MEASUREMENT SENSITIVITY, AND CALIBRATION

18. DATA COMPLETENESS

QC MEASUREMENT QUALITY OBJECTIVES

19. FIELD MEASUREMENT PRECISION

20. LAB MEASURMENT PRECISION

21. LAB MEASURMENT BIAS

22. FIELD MEASURMENT BIAS

23. BLANK CONTROL BIAS

24. ACCURACY CONTROL

STUDY DESIGN OPTIMIZATION STEPS

25. BOUNDING ERROR OR UNCERTAINTY

26. OPTIMIZING STUDY DESIGN (OPTIONAL BUT COMMONLY AN IMPORTANT STEP AFTER CONSIDERING THE FIRST 25 STEPS)

28. IMPLEMENT PILOT SCALE MONITORING

29. REVISE PLAN AND IMPLEMENT LONG-TERM MONITORING